

KERAN CHEN

Email: kchen429@wisc.edu | Tel: (+1) 6088679968 | Web: [linkedin](#)

EDUCATION

University of Wisconsin, Madison

Ph.D. in Statistics and Computer Science (Advisor: Kirthevasan Kandasamy)

Madison, U.S

2022 - 2027

University of Science and Technology of China

B.Sc. in Statistics, School of the Gifted Young

Hefei, China

2018 - 2022

RESEARCH INTERESTS

My research interests span the intersection of machine learning, mechanism design and large language models. I am particularly focused on data pricing and the development of mechanisms and marketplaces for data sharing that encourage truthful data contributions.

PUBLICATIONS AND PREPRINTS

[1] Learning to Price Homogeneous Data

Keran Chen, Joon suk Huh, Kirthevasan Kandasamy. **NeurIPS 2024**.

[2] Incentivizing Truthful Data Contributions in a Marketplace for Mean Estimation

Keran Chen, Alex Clinton, Kirthevasan Kandasamy. Preprint 2025.

[3] Gene regulatory network inference using single-cell multiome ATAC-seq and RNA-seq data

Yuge Wang, **Keran Chen**, Ting Cai, Hongyu Zhao. **American Society of Human Genetics 2022**.

RESEARCH EXPERIENCES

Learning to Price Homogeneous Data

Department of Computer Science, University of Wisconsin, Madison

Jul.2023 - May.2024

- Studied an online data pricing problem where a seller aims to maximize revenue by selling homogeneous data points to a sequence of buyers over T rounds. In each round, the seller posts an anonymous pricing curve, while each buyer strategically purchases data to maximize her own utility.
- Designed efficient discretization algorithms for approximating the revenue-optimal pricing mechanism, achieving lower time complexity with provable approximation guarantees.
- Proposed novel algorithms for the online learning version of the problem under both stochastic and adversarial settings, generalizing the upper confidence bound (UCB) and follow-the-perturbed-leader (FTPL) paradigms and establishing new sublinear regret guarantees.

Gene regulatory network inference with joint analysis of single cell ATAC-seq and RNA-seq

Department of Biostatistics, visiting student, Yale University

May.2021 - Jan.2022

- Proposed a method for constructing a gene regulatory network (GRN) based on the expression levels of human transcription factors (TF) and open chromatin accessibility.
- Developed a new method for designing the TF binding score to represent the probability that a TF binds on open chromatin, then integrated this score into our prediction model. Results showed we improved the accuracy of GRN prediction.

PROJECTS

Reinforcement Learning with low rank structure

- Showed that when the behavior policy cannot cover all the state-actions pairs, the values of uncovered state-action pairs can still be approximated if the Markov Decision Process (MDP) has a low-rank structure.
- We provided error bounds for Offline Q-learning policy evaluation in infinite-sample setting, also conducted numerical experiments to further validate the practicality of the algorithm.

Multi-Modal Temperature Prediction

- Processed time-series sensor data (temperature, humidity, pressure) with spatial features from satellite using Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). Combined these features through concatenation to capture spatial and temporal dependence.
- Implemented multi-task deep learning framework to forecast future temperature.

Exploring the generalizability of Vision Language Frontier Maps (VLFM) in real-world robotics conditions

- Built dynamic indoor environments to replicate real-world navigation conditions.
- Simulated robot vision using Chrono Sensor, enabling high-fidelity camera emulation with optical effects like lens flare, defocus flur, and distortion.
- Evaluated VLFM's performance in challenging conditions such as poor lighting, blur, and distorted optics.

Evaluation of Large Language Models on Biomedical QA Tasks

- Conducted a systematic evaluation of standard GPT-4 and our fine-tuned language model on our manually annotated biological QA dataset.
- Automated the benchmarking process with DeepEval, measuring performance across multiple dimensions including accuracy, relevance, bias, and ethical compliance.
- Automated large-scale LLM evaluation using pytest to perform batch testing and regression validation across multiple quality metrics.

AWARDS AND HONORS

Excellent Graduate, USTC	2022
Rose Endeavor Scholarship, USTC (top 10% in School of the Gifted Young)	2021
Chinese Mathematics Olympiad, Bronze	2017

SKILLS

Python, PyTorch, R, SQL, C, Matlab